Published in final edited form as:

J Proteome Res. 2015 December 04; 14(12): 4995–5006. doi:10.1021/acs.jproteome.5b00827.

Integrated Proteomic and Genomic Analysis of Gastric Cancer Patient Tissues

Julia Fangfei Yan[†], **Hoguen Kim**[‡], **Seul-Ki Jeong**[§], **Hyoung-Joo Lee**[§], **Manveen K. Sethi**[||], **Ling Y. Lee**[||], **Ronald C. Beavis**[⊥], **Hogune Im**[#], **Michael P. Snyder**[#], **Matan Hofree**[▽], **Trey Ideker**[○], **Shiaw-lin Wu**[†], **Young-Ki Paik**[‡,§], **Susan Fanayan**[◆], and **William S. Hancock**[*,†]

[†]Barnett Institute and Department of Chemistry and Chemical Biology, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115, United States

[‡]Yonsei University College of Medicine, Yonsei University, 50-1 Yonsei-Ro, Seodaemun-gu, Seoul 120-752, Korea

[§]Yonsei Proteome Research Center, Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul 120-749, Korea

[||]Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

[⊥]Department of Biochemistry and Medical Genetics, Faculty of Health Sciences, University of Manitoba, 745 Bannatyne Avenue, Winnipeg, Manitoba R3E 0J9, Canada

[#]Department of Genetics, Stanford University, Stanford, California 94305, United States

[▽]Department of Computer Science and Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, United States

[○]Program in Bioinformatics, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, United States

[◆]Department of Biomedical Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

Abstract

V-erb-b2 erythroblastic leukemia viral oncogene homologue 2, known as ERBB2, is an important oncogene in the development of certain cancers. It can form a heterodimer with other epidermal growth factor receptor family members and activate kinase-mediated downstream signaling pathways. ERBB2 gene is located on chromosome 17 and is amplified in a subset of cancers, such

[*]Corresponding Author: Phone: 1+ 617-869-8458. Fax: 1+ 617- 373-8795. wi.hancock@neu.edu.

Author Contributions

All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ADDITIONAL NOTE

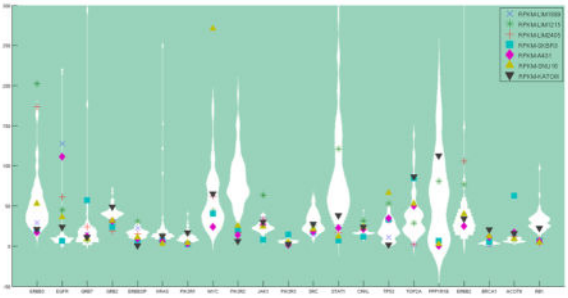
Intended as part of the The Chromosome-Centric Human Proteome Project 2015 special issue.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b00827.

as breast, gastric, and colon cancer. Of particular interest to the Chromosome-Centric Human Proteome Project (C-HPP) initiative is the amplification mechanism that typically results in overexpression of a set of genes adjacent to ERBB2, which provides evidence of a linkage between gene location and expression. In this report we studied patient samples from ERBB2-positive together with adjacent control nontumor tissues. In addition, non-ERBB2-expressing patient samples were selected as comparison to study the effect of expression of this oncogene. We detected 196 proteins in ERBB2-positive patient tumor samples that had minimal overlap (29 proteins) with the non-ERBB2 tumor samples. Interaction and pathway analysis identified extracellular signal regulated kinase (ERK) cascade and actin polymerization and actinmyosin assembly contraction as pathways of importance in ERBB2+ and ERBB2− gastric cancer samples, respectively. The raw data files are deposited at ProteomeXchange (identifier: PXD002674) as well as GPMDB.

Graphical Abstract



Keywords

ERBB2; EGFR; GRB2; RNA-Seq; gastric cancer patient tissues; Chromosome-centric Human Proteome Project

INTRODUCTION

Gastric cancers can be divided into intestinal type, diffuse type, and mixed type cancers according to the Lauren classification.[1] Carcinogenesis mechanisms as well as the stromal compositions are significantly different for the two types of gastric cancer.[1,2] In this study all of the cases are intestinal-type (GI cancer). The epidermal growth factor receptor (EGFR) superfamily of tyrosine kinase receptors,[3] of which ERBB2 is a member, is one of the important growth factor receptor systems that are commonly overexpressed in human tumors.[4] ERBB2 has been well studied in invasive breast cancer and predominantly used as biomarker for diagnosis and therapy.[5] The ERBB2 amplicon has also been shown to be significant in cancer studies, in particular, breast cancer.[5d,6] Furthermore, a subset of gastric cancer patients contain ERBB2 (Her-2) positive tumors and a monoclonal antibody directed against ERBB2 (Herceptin) is used as a treatment modality.[7,8]

In this study, we describe the proteomic analysis of two gastric cancer clinical sample sets, ERBB2-positive (ERBB2+) and the non-ERBB2-expressing (ERBB2−) tumor and matching control samples. Tissue samples were selected by immunohistochemistry staining (IHC

staining) as well as fluorescence in situ hybridization (FISH) with analysis of anti-ERBB2 reactivity.[9] Two levels of comparison are employed in our proteomic study. One comparison is between tumor tissue and its control sample in an individual patient, and the other is the comparison between ERBB2-positive and non-ERBB2-expressing samples collected from different patients. We also examined transcriptomic data from the two ERBB2-expressing gastric cancer cell lines, SNU16 and KATOIII.[7,10] We will report on the results and insights obtained by these proteomic and transcriptomic analysis.

EXPERIMENTAL SECTION

Reagents and Materials

Trypsin (sequencing grade) was purchased from Promega (Madison, WI). Guanidine hydrochloride, formic acid (FA), ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), tris, urea, thiourea, and CHAPS were purchased from Sigma-Aldrich (St. Louis, MO). Protease inhibitor cocktail tablets were purchased from Roche (Indianapolis, IN). LC–MS grade water was purchased from J. T. Baker (Philipsburg, NJ). Acetonitrile with HPLC grade was purchased from Thermo Fisher Scientific (Fairlawn, NJ). Novex sharp unstained protein standard, RPMI 1640 media and NuPAGE 4–12% Bis-Tris gels were purchased from Invitrogen (Carlsbad, CA).

Selection of Patient Samples

The patient samples were selected by Dr. Hyunki Kim from Yonsei University College of Medicine (all patient samples were obtained in Korea) based on results from immunohistochemistry staining (IHC staining) as well as FISH (fluorescence in situ hybridization). Supplemental Table S-1a describes the selection criteria in IHC staining. The ERBB2+ patient samples in this paper all have positive results in immunohistochemistry test, which has criteria of positive (3+) as “strong complete, basolateral, or lateral membranous reactivity in 10% of the tumor cells”.[11] The definition of fluorescence in situ hybridization positivity is a “HER2: chromosome 17 ratio of >2.0”. In brief, the positivity criteria was described as following “20 cohesive tumor cells showing highest gene count”, with “amplification >2.0” (“the ratio between the number of copies of the HER2 gene and the number of copies of chromosome 17 within the nucleus counted in at least 20 cancer cells were applied”).[11] The criteria and definition cites the paper “HER2 Testing in Gastric Cancer: A Practical Approach” published in *Modern Pathology* in 2012. This paper describes the details of gastric cancer sample selection criteria, which are the same criteria applied to sample selection in Korea.[11]

Therefore, a total of 150 samples were examined and eight samples were selected as ERBB2+ sample sets and nine were selected as ERBB2− sample sets. The eight ERBB2+ patient samples all have positive results in FISH and 3+ in IHC staining. (See Supplemental Figure S-1b.) On the contrary, nine ERBB2− patient samples were selected (see Supplemental Figure S-1c): Seven samples were selected based on their negative results from IHC staining. There are two samples that have positive results in IHC staining. They were retested by FISH, and both of them show negative results in FISH. These two samples were classified as ERBB2−. Upon selection, these tumor tissues and the adjacent tissues

were used for further analysis. (Control tissues are healthy tissues surrounding cancer tissues.)

Protein Extraction, Separation, Fractionation, and In-Gel Digestion

Proteins were extracted from each tumor and control tissue samples by the following steps. Each tissue sample (~50 mg) was lysed in 250 μL of lysis buffer (30 mM Tris, 7 M urea, 2 M thiourea, 65 mM DTT, 4% CHAPS, pH 8.8), containing protease inhibitor, followed by ultrasonication at 15 s intervals, 7 to 1- times, with 15 s pause between each treatment. Sections of the tissue sample, which were not readily lysed at the conclusion of the sonication process, were discarded. Each sample was then centrifuged at 14 000 rpm for 30 min at 4 °C, and supernatants, containing the extracted proteins, were carefully recovered.

The extracted proteins were separated by SDS-PAGE for proteomic analysis. Each gel lane was loaded with 40 μg proteins for the control or the tumor sample. Each gel lane was sliced into five equal fractions. Fractions were reduced by 10 mM DTT in 0.1 M ammonium bicarbonate with 30 min incubation at 56 °C, followed by 1 h of incubation with 55 mM IAA in the dark at room temperature for alkylation. For digestion by trypsin, digestion buffer was prepared as 12.5 ng/μL trypsin in 50 mM ammonium bicarbonate. Gel slices were incubated in the digestion buffer containing trypsin at 4 °C for 35 min, followed by incubation in the buffer without trypsin, at 37 °C for 14 h prior to extraction. Peptides were extracted from the gel slices by 25 mM ammonium bicarbonate and acetonitrile, followed by further extraction with 5% formic acid (37 °C, 5 min vortex), and supernatants were collected and pooled together. The digested peptide solution was concentrated into pellets for storage. The pellets were reconstituted into 10 μL of 0.1% formic acid (in water) prior to LC–MS/MS analysis. Therefore, for each control and each tumor tissue sample, five fractions were analyzed by LC–MS/MS analysis.

LC-MS Analysis by LTQ-Orbitrap

The analysis was conducted by Ultimate 3000 nano-LC (Dionex, Mountain View, CA) coupled to an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA). A self-packed C18 column (Magic, 200 Å pore, 5 μm particle size, 18 cm length, 75 μm i.d.) was used. A nanospray ion source (New Objective, Woburn, MA) was used to couple the column online to the mass spectrometer. The reconstituted 10 μL peptide solution (5 fractions per sample) was used for LC–MS/MS analysis. Five out of total 10 μL was injected for each single run. The gradient mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile. The gradient was from 2% mobile phase B for sample loading. At 2 min, mobile phase B increases to 5%. Then, a linear gradient of B solvent was applied from 5 to 50% from 5 to 85 min. Then, in the next min, mobile phase B was changed from 50 to 90%, followed by an isocratic gradient of B at 90% for 5 min. Then, the organic phase (mobile phase B) decreased back to 2% in 1 min and was maintained for another 1 min. The flow rate was maintained at 200 nL/min.

Data were obtained from LTQ-Orbitrap for LC–MS/MS analysis. The mass spectrometer was operated in data-dependent mode. Precursor ion with mass window of 400–2000 m/z was applied in full-scan MS, with mass resolution of 30 000 at m/z 400. 2+ and 3+ were

selected as the most abundant charge states. 1+ and higher charge state (4+ or higher) are also included. The full scan was followed by eight sequential LTQ (linear ion trap)-MS2 scans as MS/MS (MS2), fragmented by CID (collision induced dissociation) with 35% normalized collision energy.

Protein Identification

Peptide sequences were identified using Thermo Proteome Discoverer 1.from a human database SP.human.56.5 (Swiss-Prot, Release 56.5 of Nov 2008 contained 402 482 entries). The searching process helps to assign most probable sequences for peptides according to the fragmentation spectra using the Sequest algorithm. The spectra generated from CID-MS2 were searched against the theoretical fragmentations of the proteins under studies. The search criteria are as follows: full trypsin specificity and up to two internal missed cleavages. Precursor mass tolerance was set as 20 ppm, and fragment mass tolerance was set as 0.8 Da as initial filter of enzymatic peptides. Dynamic modifications were deamidation of asparagine, and static modification was carbamidomethylation for cysteine. False discovery rate (FDR) as 0.01 was applied.[12] Proteins with spectral counts of or more are major targets in the analysis. Experiments for proteomics were done once due to limited access to patient tissues.

The raw data files are deposited at ProteomeXchange[13] (identifier: PXD002674) at www.proteomexchange.org. The data files have also loaded onto GPMDB[14] (Global Proteome Machine Database) and can be located by searching the title of this paper "Integrated Proteomic and Genomic Analysis of Gastric Cancer Patient Tissues" at http://gpmdb.thegpm.org/thegpm-cgi/dblist_gpmnotes.pl.

Cell Line Growth Conditions and RNA-Seq Measurement

Gastric Cancer cell lines SNU16 and KATOIII were grown in RPMI 1640 medium: 37 °C/5% CO_2 containing hydrocortisone (1 mg/mL), insulin (24 U/L), and α-thioglycerol (10^{-6} M) with 10% FCS.

RNA-Seq, or RNA sequencing, has recently been applied for transcriptomic analysis.[15] RNA-Seq data were contributed by collaborators in triplicates. Nanodrop 2000 spectrophotometer was used to determine the RNA concentration.[15] Illumina HiSeq 2000 equipment was used to sequence lanes of samples, and the RNA-Seq (strand-specific) library was prepared as well to gather transcript data. RNA-Seq data are stored in Short Read Archive: SRS366582, SRS366583, SRS366584, SRS366609, SRS366610, and SRS366611. Illumina TruSeq standard procedures were followed. The library of strandspecific RNA-Seq was prepared from total RNA. The sequencing was performed as paired-end RNA-sequencing. The sequences obtained from Illumina HiSeq2000 were aligned with human genome (hg19) using Tophat, which is embedded with Bowtie and transcribed into gene transcripts (NCBI build 37.2) using Cuffinks.[15] The RPKM (reads per kilobase per million mapped reads; a measurement of gene expression represented by absolute abundance) values were also calculated based on the assembly.

RESULTS

Clinical Sample Selection

In this study we examined the proteomics of gastric cancer patients that express significant levels of ERBB2, a critical biomarker for gastric cancer, as well as a control group that also had gastric cancer but did not express the oncogene.[16] Studies using immunohistochemistry and in situ hybridization have shown that ERBB2 is overexpressed in ~7–34% gastric cancers.[16,17] The proportion of tumor type for ERBB2 positivity is as follows: ~30% of intestinal type gastric cancer, ~15% of mixed type, and ~5% of diffuse type.[11] In our study, all patient samples analyzed are of intestinal type. A total of 150 samples were examined by immunohistochemistry and FISH, and eight ERBB2-positive samples and their adjacent nontumor tissues were selected for this study. The selection criteria from immunohistochemistry are listed in Supplemental Table S-1a. Additionally, another set of gastric cancer samples and controls was collected in the same manner but were selected as non-ERBB2-expressing. In Supplementary Table S-1b,c, some of the clinical parameters of the sample selection are listed.

Proteomics Analysis of Gastric Cancer Patient Samples

After extraction of the selected tissue sample, the protein extracts were separated by 1D SDS-PAGE. Supplementary Figure S-1 shows an example of a SDS-PAGE separation of extracted proteins from one sample of gastric cancer tumor and its control. After gel analysis, each gel lane was cut into five bands and the extracted proteins were digested with trypsin and then analyzed by LC–MS/MS. Here we used high protein confidence and high peptide rank and with a false discovery rate (FDR) of <1% for protein identification.

The eight gastric cancer samples with overexpression of ERBB2, together with the matching controls, were analyzed separately in a consecutive series of analysis. In a similar manner, nine non-ERBB2 expressing (ERBB2 negative) plus control samples were analyzed as a second sample set. ERBB2 was identified with seven unique peptides in ERBB2-positive gastric cancer tumor samples (total of 32 spectral counts in all eight tumor samples) but absent in the adjacent normal tissue samples. (See Supplementary Figure S-2a for examples of the MS and MS/MS identifications of both ERBB2 and EGFR.) An assessment of the quality of the peptides observed for ERBB2 was obtained by a comparison with the ranking of the corresponding peptides in GPMDB. (See Supplementary Table S-2b.)[18] The data in GPMDB are an aggregate of observations from a large number of laboratories and sample types, so the frequency of observations is not an exact match for our study, but all of the peptides that we observed were common to many other studies.

Strategy for ERBB2-Driven Gastric Cancer Analysis

The strategy used for analysis of gastric cancer is shown in Figure 1. In-depth analysis of gastric cancer patient tissue samples provided around 3000 protein IDs, which gives a measure of the depth of protein identification. Among the identified proteins, based on the sensitivity of this analytical protocol, 196 and 159 are tumor unique proteins in the ERBB2-positive (ERBB2+) and ERBB2-negative (ERBB2−) sets, respectively. Tumor unique proteins are identified as those below the detection limit of our LC–MS analysis in matching

control tissue samples. As has been used by other investigators,[6b,15,19] we hypothesized that tumor unique proteins that are known to be interactors with known oncogenes are more likely to be of disease significance. We will discuss in the next section the selection of a panel of 21 driver oncogenes and identification of 48 interactors from the set of 196 ERBB2+ tumor unique proteins.

Combination of Proteomic Analysis with Transcriptomics

To support the proteomic study we measured the transcriptome of two gastric cell lines by RNA-Seq, as described by Snyder et al.[20] The analysis corresponded to deep reads with ~11 000 transcripts measured in each analysis. While the proteomic studies were of lesser depth (~3000 proteins identified), we and others have shown that proteomics can aid the identification of significant pathways and expression events.[6b,8,15,21]

As part of the C-HPP initiative we have been exploring the relationship between a proteomic observation and the corresponding gene location on a given chromosome.[22] It has been noted that cancer as an evolutionary process exploits regions of high transcription activity.[22] As a broad measure of the level of protein expression for a genomic region, Figure 2 shows the number of proteins identified in the ERBB2-positive proteomic study as a percentage of the number of protein coding genes on the respective chromosome. The observation of a higher percentage observed in a given chromosome can result from a combination of factors such as the constitutive expression of housekeeping proteins as well as gene amplification related to oncogene expression. An example of the high-level expression of housekeeping proteins in our study was the observation of high levels of the following chromosome 12 gene products: GAPDH, TUB1A, TUB1B, TUB1C, and KRT1. Conversely the higher % observed on chromosome 17 could be related to gene amplification related to the expression of the oncogene ERBB2. This amplification is illustrated by a copy number variation (CNV) analysis of a gastric tumor sample expressing a high level of ERBB2 and indeed does show an extensive amplification of the genomic region containing this oncogene. (See Figure 3.)

Discussion

Comparison of Tumor and Control Samples—Proteomic analysis identified around 3000 proteins in the gastric cancer tissue samples. These identifications are listed, together with their chromosome location (base start and end), band location, spectral counts, as well as the RNA-Seq values for two gastric cancer cell lines (SNU16 and KATOIII) in the Supplementary Table S-3a, and the corresponding data set for the ERBB2− sample analysis is listed in Supplementary Table S-3b. We concentrated on proteins that were detected exclusively in the tumor samples (absent in the control samples). To reduce individual variability caused by factors such as tumor grade, tumor location, as well as other patient variables, the observations for the ERBB2-positive and ERBB2-negative groups and the control samples were pooled separately (total of four pooled data sets). The numbers of proteins unique to ERBB2-positive and -negative tumors compared with their controls were 196 and 159, respectively, with only 29 proteins common between these two data sets (Figure 4). This result indicates that the classification of tumor samples into ERBB2+ and ERBB2− allowed the selection of distinct disease phenotypes.

As an example of this large data set, the 20 most abundant tumor unique proteins (out of the 196 proteins) are listed in Table 1, with different colors to denote the number of proteins which have general cancer associations (12 out of 20), including ERBB2, some are specifically gastric cancer associated (4 out of 20), as well as some structural proteins (4 out of 20).[4a,b]

ERBB2 Amplicon—This study has illustrated the effective integration of transcriptomic and proteomic data, one of the C-HPP strategies recently outlined.[5d] Another aspect of this informative approach is that the proteomic researcher can better understand the genomic context of their observations. Thus, in addition to the study of specific ERBB2 pathways as potential cancer markers, we also investigated the ERBB2 amplicon that had been previously identified in breast cancer.[6a,24] In this context, an amplicon is defined as a set of colocated genes that are also coamplified with the oncogene. In Figure 5, we plot the protein identifications in the gastric cancer study for three bands known to be part of the ERBB2 amplicon on chromosome 17, namely, q12, q21.2, and q21.3.[5d] To illustrate the genomic context of the proteomic identifications, we show the number of bases in the interval between each protein-coding gene observed in our proteomic study as well as RNA-Seq values for the cell lines. As shown in this Figure, three proximal genes, PPP1R1B, ERBB2, and GRB7, are strongly associated with the gastric cancer tumor samples.[25]

The results are consistent with a comparable study by our group in breast cancer.[6b] It has been reported[5d] that “The amplicon can be of different size in patients and may be larger in breast than other cancers. Such speculations require a better set of markers to define the size of the amplicon and the supplementation of genomic with proteomic data.” As an example of the additional information shown in Figure 5: PP1R1B was separated from ERBB2 by 4 genes whose proteins products were not observed in the proteomic study (51.5 kB) but 2 of which had RNA-Seq values greater than 1. The reason for the lack of detection of these two genes is not known but the RNA-Seq information can be used to guide future targeted studies.

A proteomics signal was observed for the genes PPP1R1B, ERBB2, GRB7, GSDMB, PSMD3, and TOP2A in the ERBB2+ but not ERBB2− sample sets. (See Figure 5.) We propose that this set of proteomic measurements could be used to define the extent of the ERBB2 amplicon in these patients. and further studies will be required to establish the generality of these observations. The comparison between ERBB2+ and ERBB2− sample sets shows the differences in protein expression of the ERBB2 amplicon. Other studies have shown that a group of genes is coamplified with the oncogene ERBB2 by such mechanism as chromosomal breakage and resynthesis.[26] Similar amplicons, which have also been reported for EGFR and Myc (c-Myc), represent a mixture of passenger genes as well as genes that are functionally linked to the oncogenic process.[27] In the case of ERBB2 the adjacent genes GRB7 (nuclear transport of ERBB2),[4b] PGAP(synthesis of the lipid raft involved in the heterodimerization of ERBB2 and EGFR),[28] and TOP2A (DNA replication and synthesis)[29] have been functionally linked, and PPP1R1B has been identified as part of an oncogenomic recombination hotspot around the PPP1R1B-ERBB2-GRB7 amplicon.[30] These observations are consistent with the suggestion that in certain cases genes with related

functions may be located in the same genomic region and exhibit coordinated expression in specific situations.[6b,31]

Pathway Analysis Based on Oncogene Interactions—Because our study was focused on the role of driver oncogenes in the gastric tumor samples that are not always measurable by proteomic studies, we combined the proteomics data from the two tumor sample sets together with RNA-Seq data for the two ERBB2-expressing gastric cancer cell lines SNU16 and KATOIII to construct Table 2. The use of cell lines for measuring the transcript data was necessitated by the limited amount of patient samples available precluded RNA-Seq analysis. In Supplementary Figure S-3, violin plots are used to summarize via histograms the values for RNA-Seq for our selected set of oncogenes obtained from a set of 58 cases of gastric cancer (TGCA_STAD) and with an overlay of the quantile normalized RPKM values for SNU16 and KATOIII as well as other cancer cell lines. Except for SKBR(RNA-Seq value for ERBB2 exceeds the range shown in the figure), which is known to express very high levels of ERBB2 transcript, the RNA-Seq data for the selected oncogenes were consistent with the range observed for the TGCA gastric cancer data set. The consistency of the data shown in this plot was supportive of the correlation of the proteomic data observed for patient samples with the RNA-Seq data. In addition, ERBB2 and EGFR, the heterodimeric partner of ERBB2 in ERBB signaling family, were observed with significant levels of transcript values in both cell lines (RPKM > 4). For SNU16 other selected oncogenes, MYC, ERBB2, GRB7, TP53, TOP2A, and GRB2 were detected with RPKM value greater than 10. For KATOIII cell line, MYC, TOP2A, and PPP1R1B have RPKM value higher than 10. The value of the transcriptomic information is shown with the observation of the high levels of MYC (RPKM values of 141 and 12) in the two cell lines. The absence of MYC in proteomic data could be either due to either technology or sample-related, but as discussed later it was helpful to include MYC in subsequent interaction studies. In addition, MYC is a well-documented partner for ERBB2 and EGFR signaling pathways.[19]

Our hypothesis in this study is that ERBB2-driven cancers have a common set of pathways, and a more detailed knowledge of individual pathways would aid in the management of this disease. Because cancer is a result of complex perturbations between different oncogene-driven pathways we examined the interaction profile of the tumor specific proteins in the ERBB2 expressing cancers.[32] Supplementary Table S-4 lists the 51 out of 199 tumor unique proteins that have interactions with oncogenes listed in Table 2. In addition, in Figure 6, we have mapped the ERBB2 interactions with other expressed oncogenes together with transcriptomic and proteomic measurements in the gastric cancer samples. Figure 6 illustrates ERBB2 and its known interactors (also oncogenes). The oncogenes that have a high degree of interaction are drawn closer to ERBB2: EGFR, ERBB3, KRAS, GRB2, GRB7, and ERBB2IP. RNA-Seq measurement is denoted by the circle size; for example, ERBB and MYC have the largest circle size. Although there are differences between the two cell lines, we applied the higher RPKM value in SNU16 or KATOIII to represent cell line RNA-Seq value. The oncogenes that have spectral counts equal or higher than have a black outline in the figure, for example, ERBB2, EGFR, SRC, GRB2, and GRB7.

Subpathway Analysis of ERBB2+ versus ERBB2− Sample Sets—On the basis of the oncogene list information, for ERBB2+ gastric cancer set, four pathways, namely, ERBB family signaling, development EGFR signaling via small GTPases, development EGFR signaling, and cytoskeleton remodeling integrin outside-in signaling pathway from the GeneGo database,[33] were examined for integration of the proteomic and transcriptomic data and are listed in Supplementary Figure S-4 (a–d). For purposes of illustration we will concentrate on a pathway that fits well with the integration of transcriptomics and proteomics data, namely, the ERBB2 signaling pathway from KEGG. (See Figure 7.)[34] Figure 7a shows the correlation between the genes listed in the KEGG pathway and the RNA-Seq values for the two gastric cancer cell lines. Transcripts for most of the genes were observed, but there are significant exceptions, such as EGF and BTC, AR, and the neuregulin genes. Also shown in the Figure are the proteomics identifications (see notation of blue stars, whereas the yellow denotes tumor unique proteins). Figure 7b shows a simplified version of the ERBB signaling pathway that lists only the subpathways with a high number of protein observations and notes identifications, which are increased in ERBB2+ versus ERBB2− patient samples. Pathvisio Version 3.2.0 was used to modify the pathway to only include genes that have RPKM value higher than 1. The MAPK signaling pathway (MEK, ERK) is clearly highlighted in this analysis, which is consistent with observation of the importance of this pathway in ERBB2+ gastric cancer and ERBB2+-driven cancers in general.[5d,6b,15,22,32]

For comparison, a subpathway associated with ERBB2− gastric cancer is shown in Figure 8, namely, actin polymerization and actinmyosin assembly contraction, which is a subpathway listed in KEGG as part of the actin cytoskeleton pathway.[34] The same analytical process as for the ERBB2+ gastric cancer samples was followed with a focus on overexpressed genes as measured by transcriptomic and proteomic measurements.[32]

CONCLUSIONS

In this work, two sets of samples, ERBB2-positive and non-ERBB2 expressing (ERBB2−) gastric cancer sample sets were analyzed. FISH experiments for anti-ERBB2 reactivity were applied for selection of ERBB2-expressing and non-ERBB2-expressing patient tissue samples. Gel loading of extracted protein samples, protein digestion, and LC–MS/MS analysis by LTQ-Orbitrap provided in-depth analysis of patient samples. Tumor samples were also compared with their adajacent nontumor tissue samples. Two gastrointestinal cancer cell lines, SNU16 and KATOIII, were used to provide integration of the RNA-Seq data with proteomics data for in-depth analysis of clinical samples. As part of the C-HPP initiative, which utilizes proteomics as well as transcriptomics data, oncogene interaction values and pathway analysis allowed the identification of pathways and associated subpathways with importance in these two subtypes of gastric cancer samples, namely, the MAPK signaling pathway and actin polymerization and actinmyosin assembly contraction for ERBB2-positive and non-ERBB2-expressing (ERBB2−) gastric cancer sample sets, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge funding support from the following research grants: (Korea) The World Class University program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10086-0) to W.S.H. and Y.-K.P. and the International Consortium Project by the Korean Ministry of Health and Welfare (to Y.-K.P., HI13C2098); (USA) NIH grant (M.P.S. and H.I.) and Grant ES014811 from National Institute of Environmental Health Sciences (NIEHS) to T.I.

References

1. Dicken BJ, Bigam DL, Cass C, Mackey JR, Joy AA, Hamilton SM. Gastric adenocarcinoma - Review and considerations for future directions. Annals of Surgery. 2005; 241:27–39. [PubMed: 15621988]
2. Yamada, T., editor. Textbook of Gastroenterology. John Wiley & Sons; Hoboken, NJ: 2011.
3. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer. 2007; 7:169–181. [PubMed: 17318210]
4. (a) UniProt Consortium. http://www.uniprot.org/(b) The Human Gene Database. http://genecards.org/(c) Yu DH, Hung MC. Overexpression of ErbB2 in cancer and ErbB2-targeting strategies. Oncogene. 2000; 19:6115–6121. [PubMed: 11156524] (d) Moelans CB, van Diest PJ, Milne ANA, Johan G, Offerhaus A. HER-2/neu Testing and Therapy in Gastroesophageal Adenocarcinoma. Pathology Research International. 2011; 2011:1.(e) Bizari L, Borim AA, Moreira Leite KR, de Toledo Goncalves F, Cury PM, Tajara EH, Silva AE. Alterations of the CCND1 and HER-2/neu (ERBB2) proteins in esophageal and gastric cancers. Cancer Genet Cytogenet. 2006; 165:41–50. [PubMed: 16490596] (f) Kawano S, Ikeda W, Kishimoto M, Ogita H, Takai Y. Silencing of ErbB3/ErbB2 Signaling by Immunoglobulin-like Necl-2. J Biol Chem. 2009; 284:23793–23805. [PubMed: 19561085] (g) Ang C, Janjigian YY, Shamseddine A, Tawil A, Lowery MA, Intlekofer A, Faraj W, Al-Olayan A, Tang L, O'Reilly EM, Geara F, Al-Kutoubi A, Kelsen DP, Abou-Alfa GK. A Case of Advanced Gastric Cancer. Gastrointest Cancer Res. 2012; 5:59–63. [PubMed: 22690259]
5. (a) Brooks JD. Translational genomics: The challenge of developing cancer biomarkers. Genome Res. 2012; 22:183–187. [PubMed: 22301132] (b) Wu WKK, Tse TTM, Sung JJY, Li ZJ, Yu L, Ch CH. Expression of ErbB Receptors and their Cognate Ligands in Gastric and Colon Cancer Cell Lines. Anticancer Res. 2009; 29:229–234. [PubMed: 19331154] (c) Kancha RK, von Bubnoff N, Bartosch N, Peschel C, Engh RA, Duyster J. Differential Sensitivity of ERBB2 Kinase Domain Mutations towards Lapatinib. PLoS One. 2011; 6:e26760. [PubMed: 22046346] (d) Paik YK, Jeong SK, Omenn GS, Uhlen M, Hanash S, Cho SY, Lee HJ, Na K, Choi EY, Yan FF, Zhang F, Zhang Y, Snyder M, Cheng Y, Chen R, Marko-Varga G, Deutsch EW, Kim H, Kwon JY, Aebersold R, Bairoch A, Taylor AD, Kim KY, Lee EY, Hochstrasser D, Legrain P, Hancock WS. The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome. Nat Biotechnol. 2012; 30:221–223. [PubMed: 22398612]
6. (a) Kauraniemi P, Barlund M, Monni O, Kallioniemi A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. Cancer Res. 2001; 61:8235–8240. [PubMed: 11719455] (b) Zhang EY, Cristofanilli M, Robertson F, Reuben JM, Mu Z, Beavis RC, Im H, Snyder M, Hofree M, Ideker T, Omenn GS, Fanayan S, Jeong SK, Paik YK, Zhang AF, Wu SL, Hancock WS. Genome wide proteomics of ERBB2 and EGFR and other oncogenic pathways in inflammatory breast cancer. J Proteome Res. 2013; 12:2805–17. [PubMed: 23647160]
7. Deng NT, Goh LK, Wang HN, Das K, Tao J, Tan IB, Zhang SL, Lee MH, Wu JN, Lim KH, Lei ZD, Goh G, Lim QY, Tan ALK, Poh DYS, Riahi S, Bell S, Shi MM, Linnartz R, Zhu F, Yeoh KG, Toh HC, Yong WP, Cheong HC, Rha SY, Boussioutas A, Grabsch H, Rozen S, Tan P. A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. Gut. 2012; 61:673–684. [PubMed: 22315472]

8. Wang YK, Gao CF, Yun T, Chen Z, Zhang XW, Lv XX, Meng NL, Zhao WZ. Assessment of ERBB2 and EGFR gene amplification and protein expression in gastric carcinoma by immunohistochemistry and fluorescence in situ hybridization. Mol Cytogenet. 2011; 4:14. [PubMed: 21689422]
9. Ginestier C, Charafe-Jauffret E, Penault-Llorca F, Geneix JN, Adelaide J, Chaffanet M, Mozziconacci MJ, Hassoun J, Viens P, Birnbaum D, Jacquemier J. Comparative multi-methodological measurement of ERBB2 status in breast cancer. J Pathol. 2004; 202:286–298. [PubMed: 14991893]
10. Dahlberg PS, Jacobson BA, Dahal G, Fink JM, Kratzke RA, Maddaus MA, Ferrin LJ. ERBB2 amplifications in esophageal adenocarcinoma. Annals of Thoracic Surgery. 2004; 78:1790–1800. [PubMed: 15511476]
11. Ruschoff J, Hanna W, Bilous M, Hofmann M, Osamura RY, Penault-Llorca F, van de Vijver M, Viale G. HER2 testing in gastric cancer: a practical approach. Mod Pathol. 2012; 25:637–50. [PubMed: 22222640]
12. Aukim-Hastie C, Garbis SD. Cell-wide Metabolic Alterations Associated With Malignancy. Methods Enzymol. 2014; 543:217. [PubMed: 24924135]
13. Proteome Xchange. http://www.proteomexchange.org
14. The GPM. http://www.thegpm.org/
15. Fanayan S, Smith JT, Lee LY, Yan F, Snyder M, Hancock WS, Nice E. Proteogenomic analysis of human colon carcinoma cell lines LIM1215, LIM1899, and LIM2405. J Proteome Res. 2013; 12:1732–42. [PubMed: 23458625]
16. Gravalos C, Jimeno A. HER2 in gastric cancer: a new prognostic factor and a novel therapeutic target. Ann Oncol. 2008; 19:1523–9. [PubMed: 18441328]
17. (a) Takehana T, Kunitomo K, Kono K, Kitahara F, Iizuka H, Matsumoto Y, Fujino MA, Ooi A. Status of c-erbB-2 in gastric adenocarcinoma: a comparative study of immunohistochemistry, fluorescence in situ hybridization and enzyme-linked immunosorbent assay. Int J Cancer. 2002; 98:833–7. [PubMed: 11948459] (b) Brien TP, Depowski PL, Sheehan CE, Ross JS, McKenna BJ. Prognostic factors in gastric cancer. Mod Pathol. 1998; 11:870–7. [PubMed: 9758367] (c) Tanner M, Hollmen M, Junttila TT, Kapanen AI, Tommola S, Soini Y, Helin H, Salo J, Joensuu H, Sihvo E, Elenius K, Isola J. Amplification of HER-2 in gastric carcinoma: association with Topoisomerase IIalpha gene amplification, intestinal type, poor prognosis and sensitivity to trastuzumab. Ann Oncol. 2005; 16:273–8. [PubMed: 15668283] (d) Hofmann M, Stoss O, Shi D, Buttner R, van de Vijver M, Kim W, Ochiai A, Ruschoff J, Henkel T. Assessment of a HER2 scoring system for gastric cancer: results from a validation study. Histopathology. 2008; 52:797–805. [PubMed: 18422971]
18. Craig R, Cortens JC, Fenyo D, Beavis RC. Using annotated peptide mass spectrum libraries for protein identification. J Proteome Res. 2006; 5:1843–1849. [PubMed: 16889405]
19. (a) Hynes NE, Lane HA. Myc and mammary cancer: Myc is a downstream effector of the ErbB2 receptor tyrosine kinase. Journal of Mammary Gland Biology and Neoplasia. 2001; 6:141–150. [PubMed: 11467449] (b) Chou YT, Lin HH, Lien YC, Wang YH, Hong CF, Kao YR, Lin SC, Chang YC, Lin SY, Chen SJ, Chen HC, Yeh SD, Wu CW. EGFR Promotes Lung Tumorigenesis by Activating miR-7 through a Ras/ERK/Myc Pathway That Targets the Ets2 Transcriptional Repressor ERF. Cancer Res. 2010; 70:8822–8831. [PubMed: 20978205]
20. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009; 10:57–63. [PubMed: 19015660]
21. Sun J, Zhang GL, Li S, Ivanov AR, Fenyo D, Lisacek F, Murthy SK, Karger BL, Brusic V. Pathway analysis and transcriptomics improve protein identification by shotgun proteomics from samples comprising small number of cells–a benchmarking study. BMC Genomics. 2014; 15(Suppl 9):S1.
22. Liu S, Im H, Bairoch A, Cristofanilli M, Chen R, Deutsch EW, Dalton S, Fenyo D, Fanayan S, Gates C, Gaudet P, Hincapie M, Hanash S, Kim H, Jeong SK, Lundberg E, Mias G, Menon R, Mu Z, Nice E, Paik YK, Uhlen M, Wells L, Wu SL, Yan F, Zhang F, Zhang Y, Snyder M, Omenn GS, Beavis RC, Hancock WS. A chromosome-centric human proteome project (C-HPP) to characterize the sets of proteins encoded in chromosome 17. J Proteome Res. 2013; 12:45–57. [PubMed: 23259914]

23. (a) Loop T, Leemans R, Stiefel U, Hermida L, Egger B, Xie FK, Primig M, Certa U, Fischbach KF, Reichert H, Hirth F. Transcriptional signature of an adult brain tumor in Drosophila. BMC Genomics. 2004; 5:24. [PubMed: 15090076] (b) Doloff JC, Waxman DJ, Jounaidi Y. Human Telomerase Reverse Transcriptase Promoter-Driven Oncolytic Adenovirus with E1B-19 kDa and E1B-55 kDa Gene Deletions. Hum Gene Ther. 2008; 19:1383–1399. [PubMed: 18771358]
24. Sircoulomb F, Bekhouche I, Finetti P, Adelaide J, Hamida A, Bonansea J, Raynaud S, Innocenti C, Charafe-Jauffret E, Tarpin C, Ayed F, Viens P, Jacquemier J, Bertucci F, Birnbaum D, Chaffanet M. Genome profiling of ERBB2-amplified breast cancers. BMC Cancer. 2010; 10:539. [PubMed: 20932292]
25. (a) Bai T, Luoh SW. GRB-7 facilitates HER-2/Neu-mediated signal transduction and tumor formation. Carcinogenesis. 2007; 29:473–479. [PubMed: 17916906] (b) Menon R, Im H, Zhang EY, Wu SL, Chen R, Snyder M, Hancock WS, Omenn GS. Distinct splice variants and pathway enrichment in the cell-line models of aggressive human breast cancer subtypes. J Proteome Res. 2014; 13:212–27. [PubMed: 24111759]
26. Moore, DD., Dowhan, D. Current Protocols in Molecular Biology. John Wiley & Sons, Inc; New York: 2002. Chapter 2: Preparation and Analysis of DNA.
27. Kim HK, Choi IJ, Kim CG, Kim HS, Oshima A, Yamada Y, Arao T, Nishio K, Michalowski A, Green JE. Three-gene predictor of clinical outcome for gastric cancer patients treated with chemotherapy. Pharmacogenomics J. 2012; 12:119–127. [PubMed: 21173787]
28. RefGene *Homo sapiens* Genes. http://refgene.com/browse/id/9606/human?p=228
29. Srikantan S, Abdelmohsen K, Lee EK, Tominaga K, Subaran SS, Kuwano Y, Kulshrestha R, Panchakshari R, Kim HH, Yang XL, Martindale JL, Marasa BS, Kim MM, Wersto RP, Indig FE, Chowdhury D, Gorospe M. Translational Control of TOP2A Influences Doxorubicin Efficacy. Mol Cell Biol. 2011; 31:3790–3801. [PubMed: 21768308]
30. Katoh M, Katoh M. Evolutionary recombination hotspot around GSDML-GSDM locus is closely linked to the oncogenomic recombination hotspot around the PPP1R1B-ERBB2-GRB7 amplicon. Int J Oncol. 2004; 24:757–63. [PubMed: 15010812]
31. (a) Claverie JM. Computational methods for the identification of differential and coordinated gene expression. Hum Mol Genet. 1999; 8:1821–1832. [PubMed: 10469833] (b) van Driel R, Fransz PF, Verschure PJ. The eukaryotic genome: a system regulated at different hierarchical levels. J Cell Sci. 2003; 116:4067–4075. [PubMed: 12972500]
32. (a) Martins-de-Souza D, Carvalho PC, Schmitt A, Junqueira M, Nogueira FC, Turck CW, Domont GB. Deciphering the human brain proteome: characterization of the anterior temporal lobe and corpus callosum as part of the Chromosome 15-centric Human Proteome Project. J Proteome Res. 2014; 13:147–57. [PubMed: 24274931] (b) Aquino PF, Lima DB, de Saldanha da Gama Fischer J, Melani RD, Nogueira FC, Chalub SR, Soares ER, Barbosa VC, Domont GB, Carvalho PC. Exploring the proteomic landscape of a gastric cancer biopsy with the shotgun imaging analyzer. J Proteome Res. 2014; 13:314–20. [PubMed: 24283986]
33. Weinmann W, Maier C, Przybylski M. Characterization of Primary Structure and Microheterogeneity of Fatty-Acid Acylated Lipoproteins by 252cf-Plasma Desorption and Electrospray Mass-Spectrometry. Fresenius' J Anal Chem. 1992; 343:63–64.
34. KEGG: Kyoto Encyclopedia of Genes and Genomes. http://www.genome.jp/kegg/

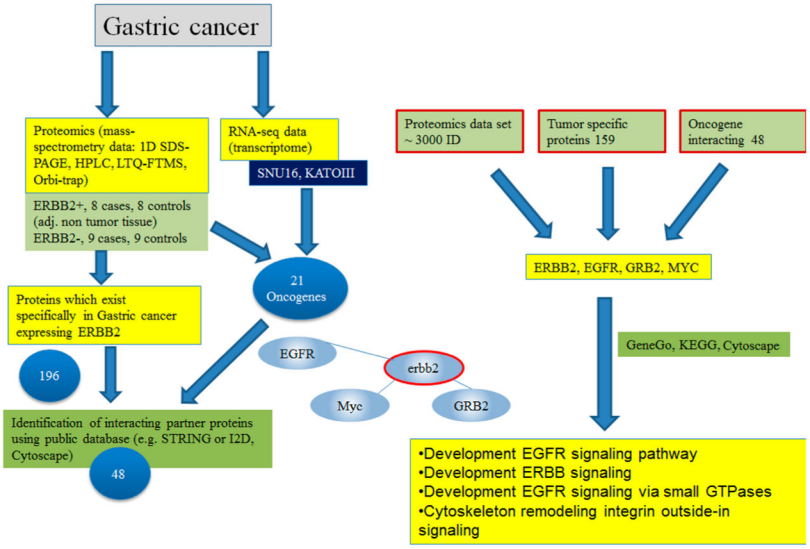


Figure 1.

Overview of the strategy for analysis of ERBB2-driven gastric cancer. Overview outlines the strategy as well as the sample set. A combination of proteomic, transcriptomic, oncogene interaction, and pathway analysis was employed in this strategy.

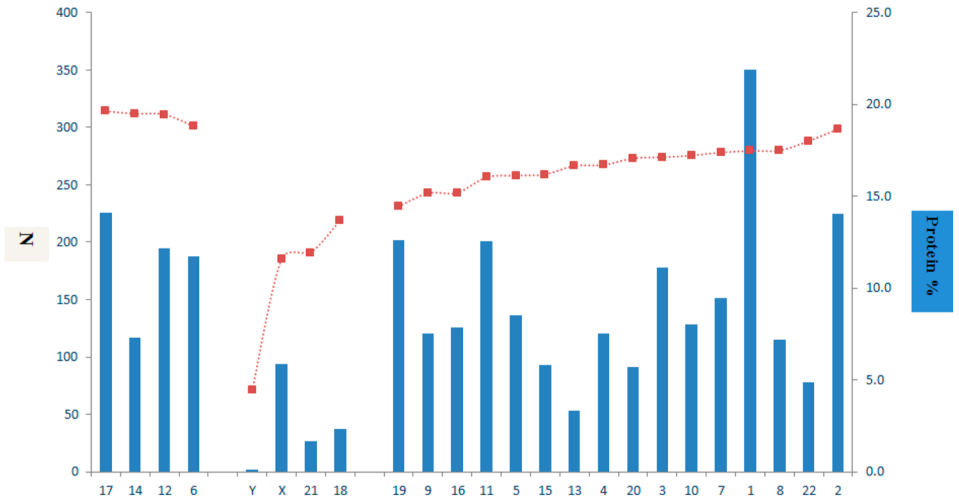


Figure 2.

Number of protein observations (N, solid bar) for a given chromosome (listed on *X* axis) and a measure of the number of protein identifications in the study of ERBB2+ gastric cancer versus the total number of protein-coding genes for the corresponding chromosome (Protein %, red squares on red dotted line; spacing is used to separate highest, lowest, and intermediate groups).

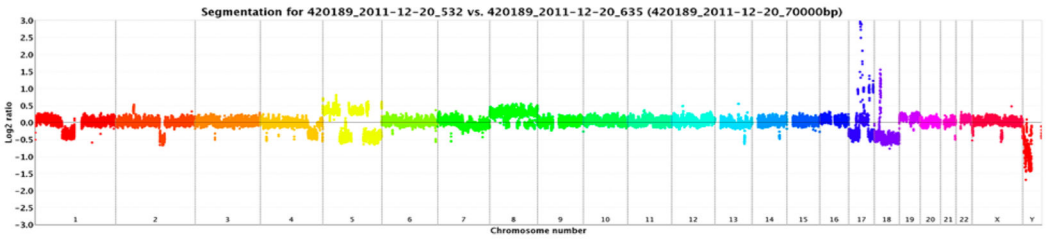


Figure 3.

CNV detection of Nimblegen human CGH 385 K for gastric cancer patient samples. Copy number variation (CNV) for each chromosome for a gastric tumor sample. The highest level of variation is in the region of the ERBB2 amplicon on chromosome 17. Test and reference cDNA were independently labeled with fluorescent dyes, cohybridized to a NimbleGen Human CGH 385 K Whole-Genome Tiling array, and scanned using a 2μm scanner. Log2-ratio values of the probe signal intensities (Cy3/Cy5) were calculated and plotted versus genomic position using Roche NimbleGen NimbleScan software. Data are displayed in Roche NimbleGen SignalMap software.

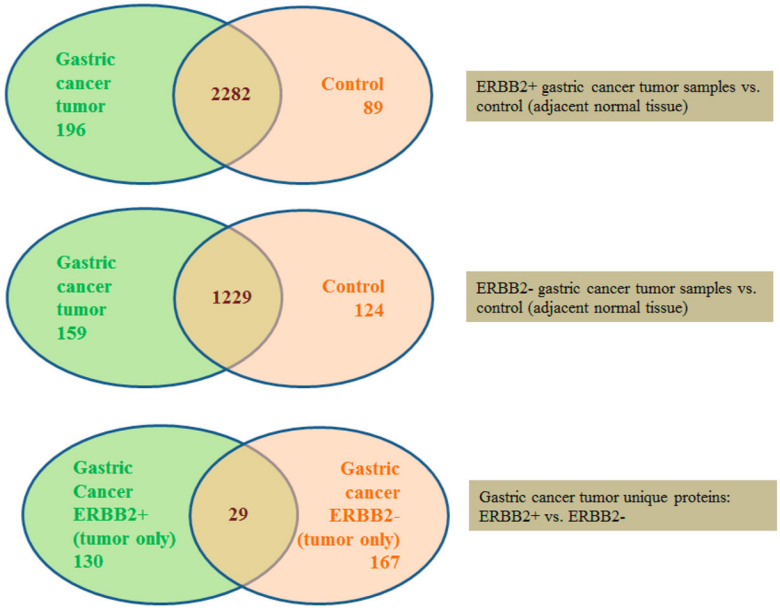


Figure 4.

Venn diagram comparing the number of unique protein identifications in ERBB2+ and ERBB2− gastric cancer and matching control samples. Figure shows the number of common (overlap region) and unique proteins in each sample set.

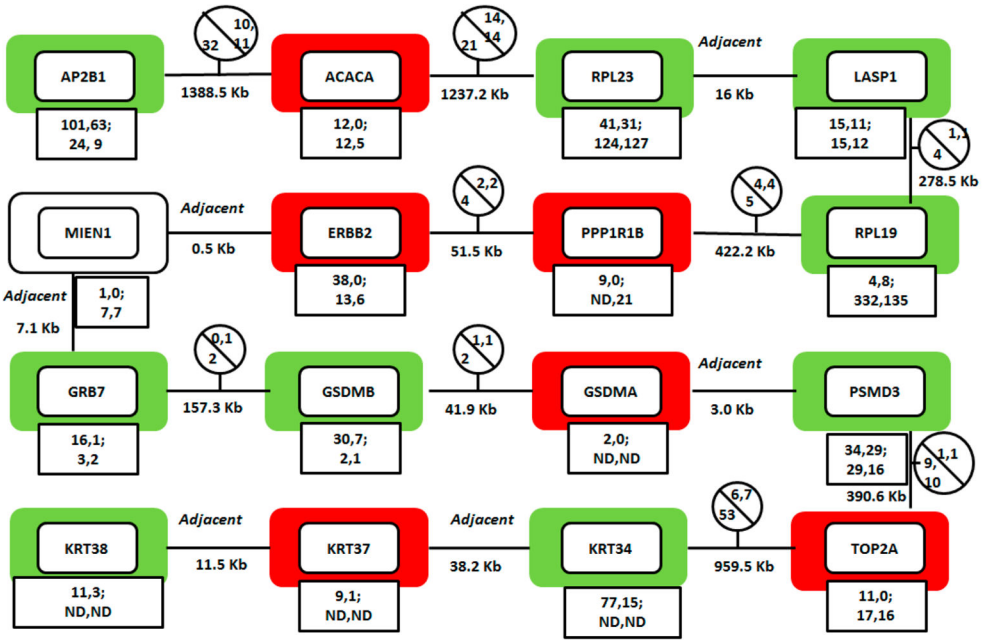


Figure 5.

Identification of genes around ERBB2 genomic region on chromosome 17 in gastric cancer and control samples. Red boxes: genes representing the proteins observed only in tumor samples of ERBB2+ sample set (not observed in control). Green boxes: genes representing the proteins observed in the both tumor and control samples of ERBB2+ sample set. White boxes: genes representing the proteins observed with spectral counts less than in all samples. MIEN1 is given a white box, as it was observed with only a single spectrum from proteomics analysis. Note that while the color of the box denotes the presence of a proteomic signal for the gastric cancer samples, the experimental data are given in the box below the gene name. Numbers in the top part: spectral counts for tumor, control; bottom part: RNA-Seq values for SNU16 and KATOIII respectively. The genomic context of each measurement is shown in the circle above the line connecting any two genes. Numbers in the left half of the black circle: number of protein-coding genes not observed in the interval between each gene observed in our proteomic study and right-hand half is the number of genes in this region with RNA-Seq values for SNU16 and KATOIII, respectively. The number of bases between the observed genes is given below the connecting line.

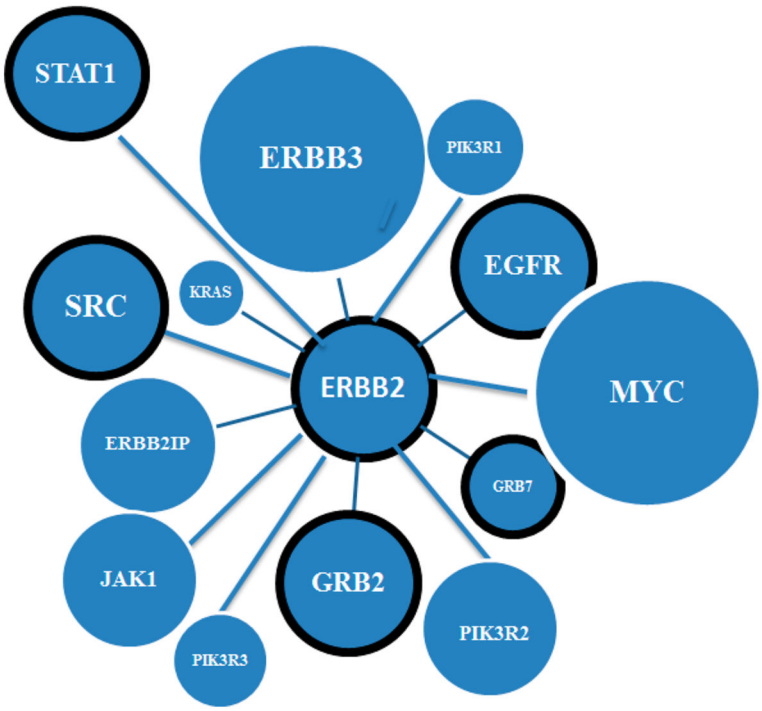


Figure 6.

Graphic illustration of the process used to prioritize oncogenes of importance for ERBB2+ via RNA-Seq, proteomics, and interaction data from GeneCards (www.Genecards.org). The following visualization assignments were used: Line length: interaction score (longer line = weaker interaction). Circle size: RPKM value (large: RPKM > 15; medium: RPKM between and 15; small: RPKM between 1 and 3; RPKM value here is denoted as the higher RPKM value in SNU16 or KATOIII for each gene). Black circle: if observed with spectral counts equal or higher than from proteomic experiments.

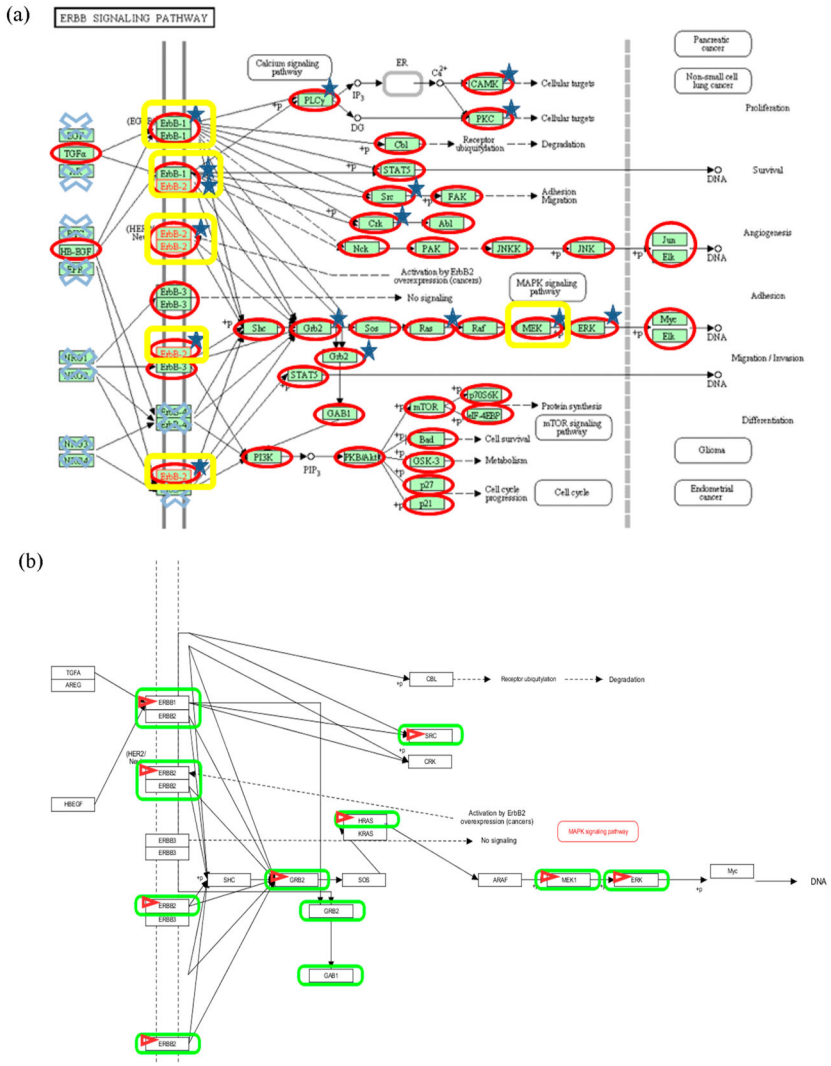


Figure 7.

ERBB signaling pathway analysis. (a) Summation of the RNA-Seq and proteomic values for genes present in the ERBB signaling pathway from KEGG.[34] Red circle: RNA-Seq ≥1 cutoff value. (The red box means the RPKM value is higher than the cut off. The red color is selected only for better clarity. Note this is a different format for visual clarity as the red circles provid best contrast with the green boxes used in the KEGG figures.) Blue cross = RNA-Seq <1. Blue star = proteomics spectral count value >0. Yellow box = proteomics spectral count value >3 only in ERBB2+ tumor sample. (b) Simplified ERBB signaling pathway that lists only the subpathways with a high number of protein observations (ERBB2+ vs ERBB2− patient samples). All genes illustrated have RPKM value >1, green box: spectral counts ≥1; the red star: higher spectral counts in ERBB2+ set (also spectral counts in tumor is higher than control). Note that color is selected only for visual clarity.

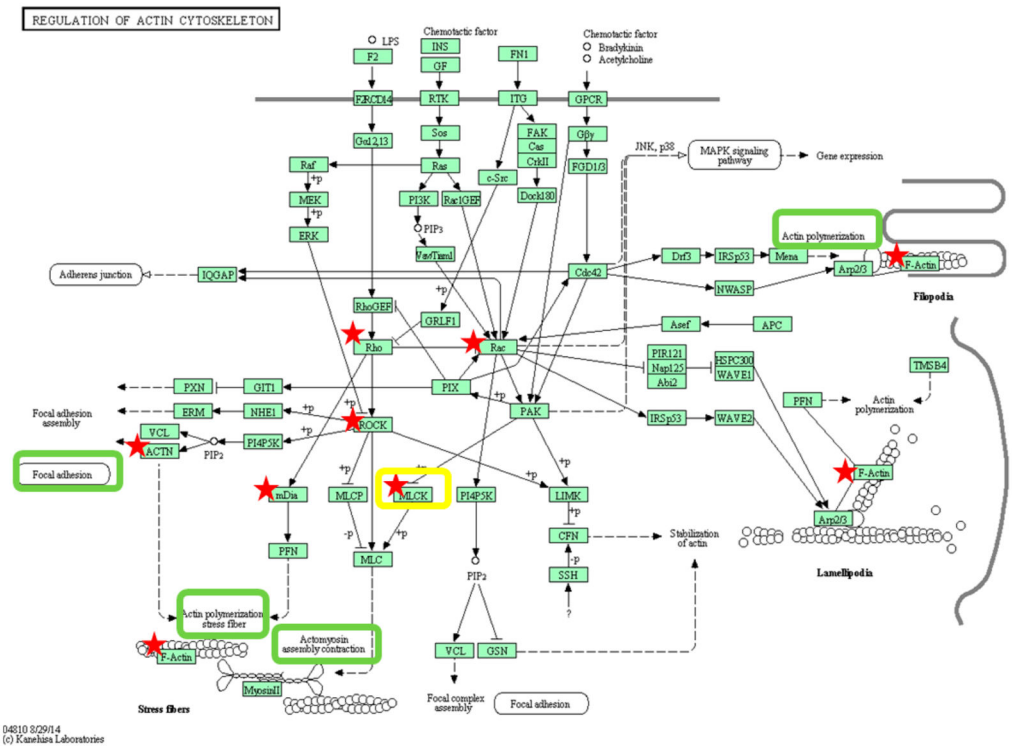


Figure 8.

Summation of the proteomic values for the ERBB2− sample set for genes present in the regulation of actin cytoskeleton pathway for ERBB2− set from KEGG. The following notations were used: Red star: proteomics value >0 in ERBB2− set (red color is only selected for better clarity in the figure). Yellow box: proteomics value >3 spectral count only in ERBB2− tumor samples. Green boxes: potentially important pathway functions. Note that this is a different format for visual clarity as the red circles provide the best contrast with the green boxes used in the KEGG figures.

Table 1

Top 20 Proteins from the 196 Tumor Unique Proteins from the ERBB2+ Gastric Cancer Sample Set

Protein name	Gene symbol	Chromosome	Gastric cancer (spectral counts)	Literature[a]
Keratin 81	KRT81	12	162	
Elongation factor (overexpression in carcinoma)	EE1A1P5	9	147	
Deydrogenase/reductase	DHRS2	14	134	
ADP-ribosylation factor (overexpression in Gastric Cancer)	ARF3	12	126	
Tubulin, beta 2B, class IIb	TUBB2B	6	71	
FK506 binding protein 10 (tumor associated)	FKBP10	17	50	
Heat shock protein 2	HSPA2	14	46	
Receptor tyrosine-protein kinase erbB-2	ERBB2	17	38	
Heterogeneous nuclear ribonucleoprotein (gastric cancer associated)	HNRNPA1	12	37	
synemin, intermediate filament protein	SYNM	15	36	
Transketolase-like (cell proliferation)	TKTL1	X	31	
defensin (expression in gastric cancer)	DEFA4	8	28	
Carboxylesterase (cancer biomarker candidate)	CESI	16	27	
versican (metastasis of cancer)	VCAN	5	26	
EGF containing fibulin-like extracellular matrix protein	EFEMP1	2	25	
Carcinoembryonic antigen-related cell adhesion molecue	CEACAM5	19	23	
tubulin, beta 8, class VIII	TUBB8	10	22	
anterior gradient homolog	ARG3	7	21	
dipeptidase 1 (cancer marker)	DPEP1	16	20	
immunoglobulin kappa variable	IGKV3–20	2	20	

[a]Green box: gastric cancer associated proteins, yellow box: cancer associated proteins, red box: structural proteins; Novoseek inferred disease relationships (GeneCards).

Table 2

List of Oncogenes in the Gastric Tumor Samples As Well As Their Interaction Scores From GeneCards[a]

gene symbol	gastric cancer spectral counts (ERBB2 expressing tumor, control; in bracket: non-ERBB2 expressing tumor, control)	RPKM value of two gastric cancer cell lines		I2D score to ERBB2	String score to ERBB2 (combined)
		SNU16	KATOIII		
ERBB3	0,0 (0,0)	17.0	3.6	7	0.999
EGFR	6,0 (0,0)	11.3	4.0	5	0.999
GRB7	16,1 (0,0)	3.0	1.8	5	0.999
GRB2	11,0 (1,5)	10.2	8.6	6	0.999
ERBB2IP	0,0 (0,0)	3.8	ND	5	0.999
KRAS	1,0 (4,1)	0.8	2.1	N/A	0.750
PIK3R1	0,0 (0,0)	1.2	2.8	5	0.996
MYC	0,0 (0,0)	141.3	11.6	N/A	0.953
PIK3R2	0,0 (0,0)	8.3	1.0	5	0.931
JAK1	0,0 (0,0)	8.0	5.1	2	0.854
PIK3R3	0,0 (0,0)	1.7	ND	2	0.825
SRC	10,1 (3,1)	7.1	4.7	6	0.999
STAT1	33,7 (19,0)	4.0	6.6	2	0.521
TP53	0,0 (0,0)	21.1	N/A	N/A	N/A
TOP2A	11,0 (0,0)	17.0	16.0	N/A	N/A
PPP1R1B	9,0 (0,0)	0.5	21.1	N/A	N/A
ERBB2	38,0 (0,0)	12.6	5.8	N/A	N/A
BRCA1	0,0 (0,0)	4.1	3.5	N/A	N/A
ACOT8	0,4 (0,0)	3.1	2.7	N/A	N/A
RB1	0,0 (0,0)	1.4	3.8	N/A	N/A

[a]Proteomics data from the two tumor sample sets (ERBB2-expressing and non-ERBB2-expressing) together with RNA-Seq data from cell lines SNU16 and KATOIII. Additional proteogenomic data as well as literature based data mining and bioinformatic databases such as STRING were obtained from Genecards (www.genecards.org).